QA
76
M94
SI- Ref

MUDPIE no. 25

Museum and University Data, Program and Information Exchange

NSF FUNDING OF COMPUTER-ORIENTED PROJECTS

The National Science Foundation has recently awarded several grants for projects of interest to MUDPIE readers, including:

$50,000 to Carnegie Institution of Washington, for development of an electronic library of rock analyses, to be made publicly available.

$3,600 to Cornell University for the development of an ecological data base for ornithological research.

$102,100 to National Oceanic and Atmospheric Administration, for development of a world data center for oceanography.

$27,550 to Queens College (CCNY), for publication of the journal Computers and the Humanities.

$75,700 to Scripps Institution for Oceanography, for preparation of a species catalog of polycystine Radiolaria.

$420,000 to the American Geological Institute, to fund the operational testing of the Geological Reference File (GeoRef), a computer-based bibliographic data bank.

$74,500 to the International Association for Plant Taxonomy to permit the continued conversion of the Index Nominum Genericarum (ING) data into machine readable form.

$46,200 to the University of Notre Dame to conclude preparation of the automated data file for the Greene Herbarium Botanical collection.

NATIONAL ENVIRONMENTAL INFORMATION SYMPOSIUM

The United States Environmental Protection Agency (EPA) has announced a symposium under the above title to be held in Cincinnati, Ohio, September 24-27, 1972, at the National Environmental Research Center of EPA in Cincinnati. The purpose of the symposium is "to provide a forum for producers and handlers of environmental data to identify and explain resources and methods in order to help users better meet their information needs; to bring together citizens' organizations, universities, libraries, professional and trade associations and governmental

AUG 22 1972

bodies to share objectives and interests related to the production, use, and dissemination of environmental information." Anyone interested should contact: Mr. G. M. Gigliotti, Director, Public Affairs Office, National Environmental Research Center, U. S. Environmental Protection Agency, Cincinnati OH 45268. Telephone: (513) 871-1820, ext. 250.

HIGH SPEED SORTING

The time required to order data by a computer is dependent upon the degree of disorder of the data and the number of records being sorted. For randomly ordered data, sorting time is proportional to the square of the number of records to be ordered when conventional sorting techniques are used. On IBM 1130 or 1800's or comparable types of computers with slow core cycle times, sorting times can run as long as 30 minutes for 500 records and 2 hours for 1000 records. Efficient sorting subroutines are often supplied by computer firms with software packages, but in many cases these subroutines are not supplied. I have developed a technique for high speed sorting which can be used on computers which do not have such a supplied routine. The technique involves initial rough sorting followed by conventional techniques, resulting in a substantial reduction of sorting time. The subroutine requires increased core storage and a longer load time, but sorting time can be reduced by a factor of 100 or greater. The subroutine has been used to reduce sorting times for 500 records from 30 minutes to 15 seconds. A Fortran IV listing for floating point data will be supplied upon request. This particular subroutine is not applicable in BASIC.--Doug Hoese, The Australian Museum, 6-8 College St., Sydney, N. S. W. 2000, Australia.

ON-LINE BIBLIOGRAPHIC SEARCHING

The National Library of Medicine, Bethesda, MD, has gone on line with a data base of more than 130,000 citations from 239 medical journals, including all material indexed since January 1, 1969. The file will be expanded to include over 1000 journals and 300,000 citations. Currently the system can be accessed through a local call in about 35 cities throughout the United States. The user must provide his own terminal, and will pay his own telephone costs. Those institutions joining the network must send a representative to Bethesda for training in the use of the system and agree to provide services to health professionals beyond the normal service level. The experimental program leading to the expanded service now offered included a data base on toxicology, so that file is already in good shape. Further information can be obtained from the National Library of Medicine, Bethesda MD 20014.--JAP.

M U S T A R D

(Museum and University STorage And Retrieval of Data)

F I R

FIR, a specialized batch processing computer system which stores and retrieves information of fish collections and literature, began capturing data in March at the National Museum of Natural Science of Canada. The hardware consists of a DCT-500 teletype, a PDP-8/L minicomputer with 4 K of memory and a mag-tape in the museum, which communicate over a phone line to a Univac 1108 computer. Input and output are on the DCT-500. Data to and from the 1108 is batched on the mag-tape unit. The 1108 updates or queries the master file and prepares output for the PDP-8/L.

Input of collection data on the DCT-500 results in output on the same device of new or corrected catalogue sheets and cards which serve as species index cards and bottle labels. Literature data can be input and results in author and subject cards production. The master file is queried via the same route and results in lists, counts, or, if desired, copies of forms. FIR offers a reasonably complete, although unrefined, ichthyological information retrieval system.

Plans are being made for a generalized time-sharing system for Canadian museums.--Don E. McAllister, National Museum of Natural Sciences, Ottawa, Ontario, K1A OM8, Canada.

RECENT LITERATURE

Colquhoun, D. Lectures on Biostatistics/An Introduction to Statistics with Applications in Biology and Medicine. Clarendon Press, Oxford, 1971: xviii + 425.

Davies, R. G. Computer Programming in Quantitative Biology. Academic Press, London, 1971: xi + 492.

Gates, D. M. Flora North America: A data bank for systematic biology. Bioscience, 21, 1971: 507.

Gonzalo Ramirez, R. C. D. & H. C. Pribor. Pattern recognition of multiple tests values as a diagnostic tool. Computers in Biology and Medicine, Vol. 2, 1972: 39-44.

Hall, A. V. Computer-based data banking for taxonomic collections. Taxon, 21, 1972: 13-25.

Hawkes, Nigel. The Computer Revolution. E. P. Dutton, 1972: 216 pp.

Kerfoot, W. C. and A. G. Kluge. Impact of the lognormal distribution on studies of phenotypic variation and evolutionary rates. Systematic Zoology, vol. 20, 1971: 459-464.

Koch, G. S., & R. F. Link. Statistical Analysis of Geological Data. Volume II. Wiley & Sons, N. Y., 1971: x + 438.

Lowry, W. K. Use of computers in information systems. Science, vol. 175, 1972: 841-846.

Patrick, E. A., K Henry & J. Altman. Computer controlled picture scanning with application to labeled biological cells. Computers in Biology and Medicine, Vol. 2, 1972: 5-14.

Patil, G. P., E. C. Pielou & W. E. Waters, Editors. Statistical Ecology, volume 1. Spatial Patterns and Statistical Distributions. Penn St Press, University Park, PA, 1971: xxviii + 582.

Pauken, R. J. and D. E. Metter. Geographic representation of morphological variation among populations of Ascaphus truei Stejneger. Systematic Zoology, vol. 20, 1971: 434-441.

Thorne, R. F. North temperate floristics and the Flora North America Project. Bioscience, 21, 1971: 1-2.

Voris, H. K. New approaches to character analysis applied to the sea snakes (Hydrophiidae). Systematic Zoology, vol. 20, 1971: 442-458.

Whiffin, T. & M. W. Bierner. A quick method for computing Wagner trees. Taxon, 21, 1972: 83-90.

August, 1972

Division of Reptiles and Amphibians
National Museum of Natural History
Washington, DC 20560